

[Home](#) / [Methodology](#) / [How data were analysed](#) / [Expected levels of performance: PILNA Proficiency Scales](#)

Expected levels of performance: PILNA Proficiency Scales

PILNA's two cognitive assessments – the numeracy assessment and the literacy (reading and writing) assessment – give scores to students.

These scores show student performance on these assessments but, alone, they do not show if a student is meeting expected levels of performance in these areas. Their scores need to be compared to expected levels of performance to create insights beyond what their performance in the PILNA assessments show.

Proficiency scales development

To create these expected levels of performance, stakeholders first needed to create scales (proficiency scales) for assessment scores to sit on in each domain (numeracy, reading, writing). Then, stakeholders needed to agree what levels on these scales year fours and year sixes should be expected to perform at.

These expected levels of performance were first constructed for PILNA 2015 to achieve two main goals:

- to provide descriptions of what students can do at various levels of performance; and
- to show results in a way that can be interpreted consistently across all participating populations.

This meant that results could be readily compared across different parts of each country's population, such as across students from urban and non-urban areas, or between girls and boys. National results can also be compared with the average achievement across the region.

A panel of experts developed and described proficiency levels, using the process summarised below.

1. A "generalised item thresholds" table was prepared, containing all items (questions) from both 2012 and 2015 cycles. This is essentially a listing of each available score point across all items, ordered by the difficulty of obtaining each score point.
2. Descriptors for each score point were attached to the ordered list. These descriptors encapsulated the key cognitive demand or the particular skill involved in obtaining each score point.
3. These descriptors were then used to develop the summary proficiency level descriptions. The 2015 items were prioritised in deciding the level cut-offs and in developing the summary level descriptions.

The set of new proficiency scale levels was developed, based on the item-to-skill mapping and placing the items on a Guttman structure (i.e. ordering the items, based on difficulty, and establishing level cut-offs, based on the skill and content grouping of the items).

Although this process results in levels that are not strictly of equal width in terms of item difficulty, the panel endeavored to make the levels as uniform as possible.

Proficiency scales for PILNA 2021

The first proficiency scales that were created were for numeracy and literacy in 2015.

PILNA 2021 has, however, split literacy into two distinct domains: reading and writing. This meant that proficiency scales for these two domains needed to be established.

It was determined that the current literacy scale could be accurately disaggregated to create a new reading proficiency scale that would be comparable to previous PILNA cycles, but this could not be done for writing. The proficiency scale for writing has not yet been developed for this reason.

The 2021 writing results will be retroactively analysed against this new writing proficiency scale once it has been established and it will be used for future PILNA cycles.

The PILNA proficiency scales use a custom scale for their scores that was informed by Item Response Theory. The ability estimates from the Item Response Theory analysis are originally reported in units that are called logits, with a mean of 0 and standard deviation of 1.

To avoid the confusion that might arise from reporting negative scores, the scaled scores that will be used for public reporting have to fit in a range that does not include negative numbers.

The ability estimates in logits were converted into a PILNA scaled score, with a mean of 500 and standard deviation of 50, using the conversion formula below, making it wide enough for current and foreseeable future needs.

$$\text{PILNA Scaled Score} = [(\text{score in logits}) \times 50] + 500$$

Summary descriptors

The summary descriptors for each proficiency level in the numeracy proficiency scale and the reading proficiency scale, and PILNA scale score that is associated with each level of performance, can be found here:

- [Numeracy proficiency level descriptors](#)
- [Reading proficiency level descriptors](#)